IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
U.S. PATENT APPLICATION FOR:

IMAGING OF A SCATTERING MEDIUM USING
THE EQUATION OF RADIATIVE TRANSFER

Inventors

Andreas H. Hielscher

Alexander D. Klose

IMAGING OF A SCATTERING MEDIUM USING
THE EQUATION OF RADIATIVE TRANSFER

5

This application claims the benefit under 35 U.S.C. §120 of prior U.S. Provisional Patent Application Serial Nos. 60/177,779 filed January 24, 2000, entitled IMAGE RECONSTRUCTION IN OPTICAL TOMOGRAPHY.

BACKGROUND OF THE INVENTION

10 **Field of the Invention**

The invention relates to the field of imaging in a scattering medium, and in particular to improved methods of reconstructing images of the spatial distribution of properties inside the scattering medium.

Background Information

15

Imaging of a scattering medium relates generally to an imaging modality for generating an image of the spatial distribution of properties (such as the absorption or scattering coefficients) inside a scattering medium through the detection of energy emerging from the medium.

A typical system for imaging based on scattered energy measures, includes at

20 least one source and detector on the surface of the medium for illuminating the medium and detecting emerging energy, respectively. The source illuminates the target medium with energy that is highly scattered in the medium so that the energy in general does not

1

travel along a straight line path through the medium, but rather, is scattered many times as it propagates through the medium. The detectors on the surface of the medium measure this scattered energy as it exits the target medium. Based on measurements of the scattered energy emerging from the target medium, a reconstructed image

5   representation of the internal properties of the medium can be generated. These systems and methods are in contrast to projection imaging systems, such as x-ray, that measure and image the attenuation or absorption of energy traveling a straight line path between a the x-ray source and a detector.

As will be appreciated, there are many instances where these techniques are

10   highly desirable. For example, one flourishing application is in the field of optical tomography. In recent years optical methods using near-infrared energy (i.e., electromagnetic radiation with wavelengths in the range of ~700-~1200 nanometers) have become increasingly important for noninvasive diagnostics in medicine. The ability to use near infrared (NIR) energy to probe the human body is of particular interest,

15   because the propagation of NIR energy in tissue is exceptionally responsive to blood oxygenation levels and blood volumes, so that blood acts as a contrast agent. These attributes permit imaging of the vasculature, and thus provide great potential for detecting cardiovascular disease, tumors, brain hemorrhages, breast cancer, rheumatoid arthritis in finger joints, and the like. Furthermore, changes in the scattering properties may also be

20   used as contrast agents. For example, rheumatoid arthritis affects the scattering coefficient of the synovial fluid, which is located inside the joints.

Systems for making measurements on human subjects using sources and detectors as described above are well known and readily available. However, a major challenge

2

remains in the development of algorithms that transform these measurements into useful images on the target medium's internal properties. Since near-infrared energy is strongly scattered in tissue, standard backprojection techniques applied in X-ray tomography have been of limited success. Various other mathematical methods have been developed for

5   solving the reconstruction problem in optical tomography. Some commonly applied schemes are modified backprojection methods, diffraction tomography, perturbation approaches, full matrix inversion techniques and elliptic system methods. Most of these methods employ model-based iterative imaging reconstruction (MOBIIR) techniques. These techniques employ a forward model that provides predictions of the detector

10   readings based on an initial guess of the spatial distribution of properties inside the medium. The predicted detector readings are then compared with experimentally measured readings using an appropriately defined objective function ϕ. The true distribution of internal properties of the target medium is then determined by iteratively updating the initial guess and performing new forward calculations with these updated

15   internal properties until the predicted data agrees with the measured data from the detector readings. The final distribution of internal properties is then displayed or printed as an image.

These known techniques, however, have several disadvantages. First, all of the known reconstruction schemes are generally based on the diffusion approximation to the

20   equation of radiative transfer. The equation of radiative transport describes the propagation of photons through a scattering medium based in part on the internal properties of the target. The diffusion approximation to this equation, makes several assumptions that reduce the complexity of the equation making it easier and faster to

3

solve. However, the diffusion approximation is only valid for example: (1) for highly scattering media where the properties of the medium μ'_s (the scattering coefficient) is much larger than μ_a (the absorption coefficient), (2) for media that do not contain strong discontinuities in optical properties, such as very low scattering and absorbing regions

5 ("void-like regions") embedded in highly scattering materials, and (3) for large source-detector pair separation, i.e., the separation between a source location and a detector location on the target medium.

If the above conditions are not met, the diffusion approximation will fail to accurately describe energy propagation through the medium. Examples of highly

10 absorbing regions in the body, where the scattering coefficient is not much larger than the absorption coefficient, are hematoma or liver tissue. Void like regions, those having low scattering and low absorbing areas, are present in areas such as the cerebrospinal fluid of the brain, the synovial fluid of human finger joints, or the amniotic fluid in the female uterus. For these cases the diffusion approximation fails and it is highly desirable to have

15 a reconstruction code that is based on the theory of radiative transfer.

Another disadvantage of the known techniques is that most of the currently available reconstruction schemes require the selection of a reference medium having properties nearly identical to the known properties of the target medium, so that there is only a small perturbation between the reference and target medium. Recently, increasing

20 attention has been paid to gradient based iterative image reconstruction (GIIR) methods, which are a subclass of MOBIIR schemes. GIIR schemes overcome the limitations of small perturbations, by using information about the gradient of the objective function (the relative difference between predicted data for the reference and measured data from the

4

target) with respect to the distribution of optical properties in a minimization scheme. However, current techniques for calculating the gradient are impractical due to the complexity of the calculation and the corresponding time required to execute the calculation.

5   For the foregoing reasons, there is a need for a reconstruction scheme that can efficiently and accurately reconstruct an image of the internal properties of a heterogeneous scattering medium, including but not limited to media that contain void-like regions and media that contain regions in which the absorption coefficient is not much smaller that the scattering coefficient.

10   **SUMMARY OF THE INVENTION**

The present method and system satisfies these needs by providing a gradient-based iterative reconstruction algorithm for imaging of scattering media using (1) the equation of radiative transfer as a forward model, and (2) an adjoint differentiation method for the fast and efficient gradient calculation used in the modification of the

15   initial guess in the updating scheme.

One embodiment of the present method and system provides a method for making an initial guess of the internal properties of a target medium, predicting the propagation of energy through the medium based on the initial guess, measuring the actual propagation of energy through the medium, updating the initial guess based on the

20   predicted data and measured data. The predicted propagation of energy is calculated using the equation of radiative transfer and an initial guess as to what the spatial properties of the target might be. The measured propagation is obtained by directing energy into the target and measuring the energy emerging from the target. An objective

<center>5</center>

function is then generated that relates the predicted values to the measured values. The objective function gives an indication of how far the initial guess of the spatial properties of the target was from the actual spatial properties of the target. A gradient of the objective function is then generated using a method of adjoint differentiation. The

5 gradient indicates how that initial guess should be adjusted to make the predicted data more closely match the measured data. This process is then repeated until the predicted and measured data are within an acceptable error. At this point the adjusted initial guess is representative of the spatial properties of the actual target and an image is generated therefrom.

10 The foregoing specific objects and advantages of the invention are illustrative of those which can be achieved by the present invention and are not intended to be exhaustive or limiting of the possible advantages that can be realized. Thus, the objects and advantages of this invention will be apparent from the description herein or can be learned from practicing the invention, both as embodied herein or as modified in view of

15 any variations which may be apparent to those skilled in the art. Accordingly, the present invention resides in the novel parts, constructions, arrangements, combinations and improvements herein shown and described.

BRIEF DESCRIPTION OF THE DRAWING

For a better understanding of the invention, together with the various features and

20 advantages thereof, reference should be made to the following detailed description of the preferred embodiments and to the accompanying drawings wherein:

FIG. 1 is a flow chart of the reconstruction process;

FIG. 2 is a schematic of an exemplary measurement and imaging system;

FIG. 3A is a computational graph of the forward model calculating the objective function; and

FIG. 3B is a computational graph of the adjoint differentiation technique.

DETAILED DESCRIPTION OF THE INVENTION

5   I.   **Introduction**

As discussed above, imaging of a scattering medium refers generally to the methods and techniques of reconstructing an image of the internal properties of a scattering medium based on the transmission of energy into the medium and the measurement of the scattered energy emerging from the medium.

10   The present method and system employs an iterative image reconstruction scheme that has three major elements. The following sections describe in detail how the three major elements of the present method and system work together to yield a reconstructed image of the internal properties of the scattering target medium. First, the solution of the forward model of the equation of radiative transfer is solved using an upwind difference

15   scheme, even-parity finite-element formulation or the like; by way of example, the present invention is illustrated using the upwind scheme. This is followed with a detailed description of the second and third component, namely the objective function and the updating scheme, in which adjoint differentiation is used for the gradient calculation. Referring to FIG. 1, these elements are illustrated as the forward model 105, the objective

20   function/analysis scheme 120 and the updating scheme 125.

The forward model starts at step 100 with an initial guess $\mu_0 = [\mu_{s.0}(r), \mu_{a.0}(r)]$ (scattering and absorption coefficients respectively) of the internal properties, known

7

energy source $S(r_s)$ at the positions r_s and given boundary conditions. Based on the initial

guess μ_0, energy source $S(r_s)$ and the given boundary conditions the forward model (1)

gives a numerical solution of the energy distribution $\Psi(r)$ inside the target scattering

medium, and (2) predicts the energy radiance Ψ_d on the boundary Ω of the medium based

5   on the equation of radiative transfer. Referring to step 110 of FIG. 1, the vector **P**,

generated by the forward model, contains the predicted radiance value Ψ_d, or derived

parameter related thereto, at each detector, for each source detector pair, as a function of

the properties μ. Measured data for the actual target to be imaged is then collected in

step 115. The measured radiance at each detector, for each source detector pair is stored

10   in the vector **M**. A typical imaging system for collecting the measured data is presented

in detail below.

In the analysis scheme, step 120, the predicted radiances $\Psi_d(\mu_0)$ are compared

with the set of measured radiances **M** on the boundary Ω of an actual target medium to

define an objective function Φ. In steps 125 through 135, an optimization technique is

15   used to minimize the objective function Φ, e.g., the normalized difference between the

predicted and measured values. For this optimization the internal properties μ_0 are

updated using the derivative $d\Phi/d\mu$ of the objective function with respect to the internal

properties. In step 130, the gradient is computed by an adjoint differentiation technique,

that takes advantage of a well-chosen numerical discretization of the forward model.

20   Once the gradient is computed, the initial guess is modified in step 140 or 145,

depending on whether an inner or outer iteration is being performed, and a new forward

calculation is performed based on the new set of internal properties $\mu_0 + \Delta\mu$. The

8

iteration process continues until the minimum of the objective function is reached within a specified error. At this point the predicted detector readings are identical to the measured detector readings within a given tolerance. The internal properties, μ, of the target medium are then mapped into an image.

5   For illustration purposes, the present system and method is described in further detail below with respect to an optical tomography system. However, it will be appreciated by those skilled in the art that the methodology of the present invention is applicable for any energy source (e.g., electromagnetic, acoustic, etc.), any scattering medium (e.g., body tissues, oceans, foggy atmospheres, geological strata, and various

10   materials, etc.), and any source condition (e.g., time-independent, time-harmonic, time-resolved). Its applicability is dependent only on the presence of the phenomenology described herein, (i.e., radiative transfer being the principal mechanism of energy transport), which is expected in all cases where scattering occurs. Accordingly, this methodology can be extended to allow for new imaging approaches in a broad range of

15   applications, including nondestructive testing, geophysical imaging, medical imaging, and surveillance technologies.

II.   Exemplary Measuring System

There are many known imaging systems for collecting the measured data used in image reconstruction. A schematic illustration of an exemplary optical tomography

20   system is shown in FIG. 2. This system includes a computer 202, energy source 204, a fiber switcher 208, an imaging head 210, detectors 212, a data acquisition board 214, source fibers 220 and detector fibers 222.

9

The energy source 204 provides optical energy, directed through the fiber switcher 208 to each of the plurality of source fibers 120 one at a time in series. The source fibers 220 are arranged around an imaging head 210 so that the energy is directed into the target medium 216 at a plurality of source locations around the target.

5   The energy leaves the source fiber 220 at the imaging head 210 and enters the target medium 216 centered in the imaging head 210. The energy is scattered as it propagates through the target medium, emerging from the target medium at a plurality of locations. The emerging energy is collected by the detector fibers 222 arranged around the imaging head 210. The detected energy then travels through the detector fibers 222 to

10   detectors 212 having energy measuring devices that generate a signal corresponding to the measurement. The data acquisition board 214 receives the measurement signal for delivery to the computer 202.

This process is repeated for each source position so that a vector of measures are obtained for all of the detectors and source locations. This vector of measured data is the

15   vector **M** referred to above. The vector of data **M** is stored by the computer 202 for use in image reconstruction and other analysis discussed below.

III.   Forward Model

As discussed in the introduction above, the present method and system compares measured data **M** from the actual target with predicted data **P** for the target. The

20   predicted data is obtained through the "forward" model.

In the present method and system, the equation of radiative transfer is used for the forward model. Referring to FIG. 1, steps 100 through 110, the radiative transfer equation is used to predict the detector readings of energy emerging at one or more

10

detector locations on the scattering medium based on an initial guess of the properties of the medium. The equation of radiative transfer is derived by considering energy conservation in a small volume. The equation for time-dependent case can be written as:

$$\frac{1}{c}\frac{\partial \Psi(\mathbf{r},\omega,t)}{\partial t} = S(\mathbf{r},\omega,t) - \omega \cdot \nabla\Psi(\mathbf{r},\omega,t) - (\mu_a + \mu_s)\Psi(\mathbf{r},\omega,t) + \mu_s \int_0^{2\pi} p(\omega,\omega')\Psi(\mathbf{r},\omega',t)d\omega'. \quad (1a)$$

5   and; for the time-independent case can be written as:

$$\omega\nabla\Psi(\mathbf{r},\omega) + (\mu_a + \mu_s)\Psi(\mathbf{r},\omega) = S(\mathbf{r},\omega) + \mu_s \int_0^{2\pi} p(\omega,\omega')\Psi(\mathbf{r},\omega')d\omega', \quad (1b)$$

where $\mathbf{r}$ is the spatial position vector, ω is a unit vector pointing in the direction of photon propagation, $\Psi(\mathbf{r},\omega,t)$ and $\Psi(\mathbf{r},\omega)$ are the energy radiance in units of W cm^{-2} sr^{-1}. $S(\mathbf{r},\omega,t)$ and $S(\mathbf{r},\omega)$ are the source terms representing energy directed into a solid angle centered on ω in a unit volume at $\mathbf{r}$, μ_s is the scattering coefficient given in units of cm^{-1}, μ_a is the absorption coefficient given in units of cm^{-1} and $p(\omega,\omega')$ is the phase function that describes the probability that a photon with direction ω' will be scattered into the direction ω during a scattering event.

The scattering phase function $p(\omega,\omega')$ can be described, for example, using the well known Henyey-Greenstein scattering function,

$$p(\cos\theta) = \frac{1-g^2}{2(1+g^2-2g\cos\theta)^{3/2}}. \quad (2)$$

where θ is the angle between the two directions ω and ω'. The parameter g is the anisotropy factor which is commonly used to characterizes the angular distribution of scattering. It can range from g=-1 (total backward scattering), over g=0 (isotropic scattering), to g=1 (strictly forward scattering). Biological tissues typically have a g-

597699_1

factor between g=0.7 and g=0.98. Other scattering phase functions are also possible, for example derived from Mie theory.

The integral of the radiance over all angles ω at one point r yields the fluence rate $\Phi(\mathbf{r})$ for the time-independent case:

$$\Phi(\mathbf{r}) = \int_{2\pi} \Psi(\mathbf{r},\omega)\, d\omega \qquad\qquad (3)$$

5

While it will be appreciated to those skilled in the art that a similar three-dimensional and/or time dependent equation can be generated, by way of example the remainder of the specification will focus on the two-dimensional, time-independent problem.

10   Various computational techniques exist, such as those disclosed by Lewis E E, Miller W F, Computational Methods of neutron transport, New York, John Wiley & Sons, 1984, that numerically solve the equation of radiative transfer. Techniques commonly applied include the singular eigenfunction method, the method of spherical harmonics, the method of characteristics, the finite-element method, and the finite-

15   difference discrete-ordinate method. A concise review of these techniques has been presented by Sanchez R, McCormick N J, A Review of neutron transport approximations. Nucl. Sci. Eng. 1982,80:481-535; and McCormick NJ, Inverse radiative transfer problems: a review, Nucl. Sci. Eng. 1992,112:185-198. The discrete-ordinates method is widely used with several different finite-difference approximations, such as the diamond-

20   difference scheme, the weighted diamond-difference scheme, or the centered-difference scheme.

Another means of solving the equation of radiative transfer is the upwind-difference scheme used in connection with the discrete-ordinates method to the equation

12

of radiative transfer disclosed by Sewell G., The numerical solution of ordinary and partial differential equations, San Diego, Academic Press, 1988. This method is a computationally efficient method for the calculation of the radiance and has the advantage that it easily supplies the required mathematical structure for the adjoint

5   differentiation calculation. The adjoint differentiation method is crucial for obtaining the gradient of the objective function in an computationally efficient way, and thus obtaining an update of the initial guess of the properties of the target medium. The adjoint differentiation method and the objective function will be described in detail below in Section IV.

10   Returning to the forward problem, to solve the equation of radiative transfer with an upwind-difference discrete-ordinates method, the angular and spatial variables have to be discretized. First, the integral term in equation 1b is replaced by a quadrature formula that uses a finite set of K angular directions ω_k with $k = 1, \ldots K$. This yields a set of K coupled ordinary differential equations for the angular-dependent radiance

15    $\Psi_k(\mathbf{r}) = \Psi(\mathbf{r}, \omega_k)$ in the directions ω_k. The coupling term is the internal source term

$\mu_s \sum\limits_{k'=1}^{K} a_{k'} p(\omega_{k'}, \omega_k) \Psi(\mathbf{r}, \omega')$. The parameter $a_{k'}$ is a weighting factor that depends on the chosen quadrature formula. In this work the extended trapezoidal rule is employed.

$$\omega_k \nabla \psi_k(r) + (\mu_a + \mu_s)\psi_k(r) = S_k(r) + \mu_s \sum_{k'=1}^{k} a_{k'} p_{kk'} \psi_{k'}(r) \quad (4)$$

Additionally, the spatial variable $\mathbf{r}$ needs to be discretized. The domain Ω is

20   defined by a rectangular spatial mesh with M grid points on the x-coordinate and N grid points on the y-coordinate. The distance between two adjacent grid points along the x-

13

axis is Δx and along the y-axis is Δy. The angular radiance at the grid point (i,j) with position $\mathbf{r} = (x_i, y_j)$ for a particular direction ω_k is represented by $\Psi_{k,i,j} = \Psi_k(x_i, y_j)$. The angular direction ω_k can be expressed in cartesian coordinates with $\xi_k = \mathbf{e}_x \cdot \omega_k = \cos(\omega_k)$ and $\eta_k = \mathbf{e}_y \cdot \omega_k = \sin(\omega_k)$. The transport equation can now be written as:

5
$$\xi_k \frac{\partial \psi_k(x_i, y_j)}{\partial x} + n_k \frac{\partial \psi_k(x_i, y_j)}{\partial y} + (\mu_a + \mu_s)\psi_k(x_i, y_j) = S_k(x_i, y_j) + \mu_s \sum_{K'=1}^{K} {}^a K' p_{kk'} \psi_{k'}(x_i, y_j) \quad (5)$$

Finally, the spatial derivatives have to be replaced with a finite-difference scheme. The upwind-difference formula depends on the direction ω_k of the angular-dependent radiance Ψ_k. Thus, the set of all angular directions ω_k are subdivided into four quadrants to generate four different difference formulas for the radiance $\Psi_{k,i,j}$, depending

10   on the sign of ξ_k and η_k :

$$\text{I) } \xi_k > 0, \eta_k > 0: \quad \frac{\partial \Psi}{\partial x} \approx \delta_x \Psi_{k,I,J} = \frac{\Psi_{I,J} - \Psi_{I-1,J}}{\Delta x}, \frac{\partial \Psi}{\partial y} \approx \delta_y \Psi_{k,I,J} = \frac{\Psi_{I,J} - \Psi_{I,J-1}}{\Delta y} \quad (6a)$$

$$\text{II) } \xi_k < 0, \eta_k > 0: \quad \frac{\partial \Psi}{\partial x} \approx \delta_x \Psi_{k,I,J} = \frac{\Psi_{I+1,J} - \Psi_{I,J}}{\Delta x}, \frac{\partial \Psi}{\partial y} \approx \delta_y \Psi_{k,I,J} = \frac{\Psi_{I,J} - \Psi_{I,J-1}}{\Delta y} \quad (6b)$$

$$\text{III) } \xi_k > 0, \eta_k < 0: \quad \frac{\partial \Psi}{\partial x} \approx \delta_x \Psi_{k,I,J} = \frac{\Psi_{I,J} - \Psi_{I-1,J}}{\Delta x}, \frac{\partial \Psi}{\partial y} \approx \delta_y \Psi_{k,I,J} = \frac{\Psi_{I,J+1} - \Psi_{I,J}}{\Delta y} \quad (6c)$$

15
$$\text{IV) } \xi_k < 0, \eta_k < 0: \quad \frac{\partial \Psi}{\partial x} \approx \delta_x \Psi_{k,I,J} = \frac{\Psi_{I+1,J} - \Psi_{I,J}}{\Delta x}, \frac{\partial \Psi}{\partial y} \approx \delta_y \Psi_{k,I,J} = \frac{\Psi_{I,J+1} - \Psi_{I,J}}{\Delta y} \quad (6d)$$

The time-independent radiative transfer equation, with the external and internal source terms on the right-hand-side, can now be written as:

$$\xi_k \cdot \delta_x \Psi_{k,I,J} + \eta_k \cdot \delta_y \Psi_{k,I,J} + (\mu_a + \mu_s)\Psi_{k,I,J} = S_{k,I,J} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'} \Psi_{k',I,J} \quad (7)$$

14

Recasting the left-hand side of the preceding equation as a single operator acting

upon $\Psi_{k,i,j}$, produces;

$$\left\{\xi_k \cdot \delta_x + \eta_k \cdot \delta_y + \left(\mu_a + \mu_s\right)\right\}\Psi_{k,i,j} = S_{k,i,j} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'} \Psi_{k',i,j}, \qquad (8)$$

from which it is apparent that the system of equations 8 corresponding to all K

5   directions can be written as a single matrix equation:

$$A\,\Psi = b; \qquad (9)$$

proportional to one row of the matrix A and with an ordering of the radiance vector for a

fixed k with, for example $\xi_k > 0, \eta_k > 0$, $\Psi = (\Psi_{k,1,1}, \Psi_{k,1,2}, \dots, \Psi_{k,2,1}, \Psi_{k,2,2}, \dots, \Psi_{k,i,j},$

$\dots, \Psi_{k,M,N})$ produces:

10     $$\xi_k \frac{\Psi_{k,i,j} - \Psi_{k,i-1,j}}{\Delta x} + \eta_k \frac{\Psi_{k,i,j} - \Psi_{k,i,j-1}}{\Delta y} + \left(\mu_a + \mu_s\right)\Psi_{k,i,j} = S_{k,i,j} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'} \Psi_{k',i,j} \quad (10)$$

The resulting system of equations for the radiance $\Psi_{k,i,i}$ for all grid points is

solved for each ordinate index k by a Gauss-Seidel method. Accordingly, the matrix A is

split into a diagonal part **D**, an upper-triangular part **U**, and a lower-triangular part **L**,

with $A = L + D + U$. The original matrix equation 9 is now be written as either:

15     $$(L + D + U)\,\Psi = b \qquad (11a)$$

or, $$(D + L)\,\Psi = -U\,\Psi + b \qquad (11b)$$

In this case an upper matrix U does not exist. Thus, for example for the case

$\xi_k > 0$, $\eta_k > 0$:

$$\left\{\frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + \left(\mu_a + \mu_s\right)\right\}\Psi_{k,i,j} - \frac{\xi_k}{\Delta x}\Psi_{k,i-1,j} - \frac{\eta_k}{\Delta y}\Psi_{k,i,j-1} = S_{k,i,j} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'} \Psi_{k',i,j} \quad (12)$$

20   diagonal matrix D   lower matrix L   b

15

Now the iterative form with the iteration matrix (**D**+**L**) is expressed as

$$(\mathbf{D} + \mathbf{L})\Psi^{z+1} = -\mathbf{U}\Psi^{z} + b, \qquad (13)$$

and for the case $\xi_k > 0$, $\eta_k > 0$:

$$5 \quad \left\{ \frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s) \right\} \Psi^{z+1}_{k,i,j} - \frac{\xi_k}{\Delta x}\Psi^{z+1}_{k,i-1,j} - \frac{\eta_k}{\Delta y}\Psi^{z+1}_{k,i,j-1} = S_{k,i,j} + \mu_s \sum_{k'=1}^{K} a_{k'}p_{k,k'}\Psi^{z}_{k',i,j}, \quad (14)$$

$$\Psi^{z+1}_{k,i,j} = \frac{\left\{ S_{k,i,j} + \mu_s \sum_{k'=1}^{K} a_{k'}p_{k,k'}\Psi^{z}_{k',i,j} + \frac{\xi_k}{\Delta x}\Psi^{z+1}_{k,i-1,j} + \frac{\eta_k}{\Delta y}\Psi^{z+1}_{k,i,j-1} \right\}}{\left\{ \frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s) \right\}}. \qquad (15)$$

All $\Psi^{z+1}_{k,i-1,j}$ and $\Psi^{z+1}_{k,i,j-1}$ are already calculated at the current iteration step because of the vector ordering. For all other ordinates ω_k besides $\xi_k > 0$, $\eta_k > 0$, the radiance vector Ψ has to be re-ordered to get the same matrix structure with

10 $(\mathbf{D} + \mathbf{L})\Psi = -\mathbf{U}\Psi + \mathbf{b}$. The iteration process is repeated until the error norm $E = \left\| \Psi^{z+1}_{k,i,j} - \Psi^{z}_{k,i,j} \right\|$ at the grid point (i,j) is smaller than a defined ε.

A significant improvement in convergence speed can be achieved by a slight modification to the Gauss-Seidel method. The successive over-relaxation (SOR) method uses a relaxation parameter ρ with $1 < \rho < 2$ in order to correct the solution $\Psi^{z+1}_{k,i,j}$ of the

15 Gauss-Seidel iteration, now denoted as $\overline{\Psi}^{z+1}_{k,i,j}$. The updated value $\Psi^{z+1}_{k,i,j}$ of the SOR is a linear combination of the Gauss-Seidel value $\overline{\Psi}^{z+1}_{k,i,j}$ and the previously computed value $\Psi^{z}_{k,i,j}$ of the SOR using:

$$\Psi^{z+1}_{k,i,j} = (1 - \rho)\Psi^{z}_{k,i,j} + \rho\overline{\Psi}^{z+1}_{k,i,j}. \qquad (16)$$

597699_1

The boundary conditions are treated between each successive iteration steps. Because of the refraction index mismatch at the air-tissue interface, the outgoing radiance is partly reflected on the boundary and only a fraction of that energy escapes the medium. The internally reflected energy contributes further to the photon propagation inside the medium, while the transmitted energy enters the detectors. Using Fresnel's formula, the transmissivity T and reflectivity R are calculated at the boundary grid points for each ordinate ω_k and for the given refractive indices. The reflectivity R and the transmissivity T are defined as

$$R = \frac{\sin^2(\omega_{trans} - \omega_{in}) + \sin^2(\omega_{trans} + \omega_{in})}{2}. \tag{17}$$

$$T = 1 - R \tag{18}$$

The angles ω_{trans}, which pertain to radiances escaping the object, are determined by Snell's law ($n_{object} \sin\omega_{in} = n_{air} \sin\omega_{trans}$) given the angle ω_{in} of the radiance, which hits the boundary inside the object. The angle $\omega_{ref} = -\omega_{in}$ is just the angle of the reflected radiance on the boundary inside the object. The transmitted and the reflected radiance are calculated with:

$$\Psi_{trans}^z = T \cdot \Psi_{in}^z, \tag{19}$$

$$\Psi_{ref}^z = R \cdot \Psi_{in}^z. \tag{20}$$

The fluence Φ_{ij} on the boundary at the grid point (i,j), which enters the detector, is calculated for a given detector aperture AP using the transmitted radiance,

$$\Phi(x_i, y_i) = \int_{AP} \Psi(x_i, y_i, \omega) d\omega \approx \sum_{k=k_1}^{k=k_2} a_k \Psi_{k,i,j} = \Phi_{i,j}. \tag{21}$$

17

The weighting factors a_k are given by the extended trapezoidal rule, [84] which provides the quadrature formula in this work.

These fluence values Φ_{ij} on the boundary provide the vector of predicted values shown in step 110 of FIG. 1 used for the reconstruction algorithm.

5   IV.   **Objective Function/Analysis Scheme and Updating Scheme**

The first components of the present method and system, i.e., the forward model used to generate the predicted energy emerging from the target, and the collection of measured energy emerging from the target, have been described in detail above. The following sections describe the second and third components of the method and system,

10   illustrated in steps 120 through 135 of FIG. 1, namely, the objective function and the updating scheme, in which adjoint differentiation is used for the gradient calculation. These second and third components are used to modify the initial guess of the internal properties of the medium in an iterative process.

1.   **Objective Function/Analysis Scheme**

15   Optical tomography can be viewed as an optimization problem with a non-linear objective function. Referring to FIG. 1, step 120, the discrepancy between the actual measurements, represented by the vector **M**, and the predictions, given by the vector **P**, for m source-detector pairs is defined by a scalar-valued objective function $\Phi(P)$. The predictions **P** are mapped onto a scalar φ using the χ^2 - error norm:

20
$$\Phi : \Re^m \to \Re \qquad (22a)$$

$$P \mapsto \varphi = \frac{1}{2} \sum_i^m \left(P_i - M_i\right)^2 \qquad (22b)$$

18

The predictions **P** are determined for all m source-detector positions by evaluating

the forward model F, given the spatial distribution of internal properties μ (see Part I):

$$F : \mathfrak{R}^n \to \mathfrak{R}^m \qquad (23a)$$

$$\mu \mapsto P(\mu) \qquad (23b)$$

5   The vector μ is of length $n = 2\,I\,J$, and contains the internal properties μ_s and μ_a.

The parameters I and J denote the number of grid points of a finite-difference grid along

the x-axis and y-axis, respectively.

The goal in image reconstruction is to determine a vector μ that minimizes the

objective function Φ. As already mentioned in the introduction, gradient-based

10   optimization techniques, such as the conjugate-gradient method, employ the gradient

$\nabla_\mu\Phi(\mu)$ to calculate a sequence of points μ_0, μ_1, ..., μ_i that lead to ever-improving

values of Φ. This iterative process is terminated when $|\Phi(\mu_{i+1}) - \Phi(\mu_i)|$ becomes smaller

than a predefined value ε. A crucial task within this process is the computationally

efficient calculation of the gradient $\nabla_\mu\Phi(\mu_i)$. The following section describes in detail

15   how this can be done using an adjoint differentiation scheme.

2.   Gradient Calculation With Adjoint Differentiation

In adjoint differentiation schemes the numerical code that calculates the objective

function Φ is directly differentiated to obtain the gradient $\nabla_\mu\Phi(\mu_i)$, FIG. 1, step 130. To

apply this method one has to decompose the objective function into a series of elementary

20   differentiable function steps. By systematically applying the chain rule of differentiation

to each of these elementary steps in the reverse direction of the forward model

computation, a value for the gradient is obtained.

19

The parameter $\Phi = \Phi(P(\mu))$ can be decomposed into sub-functions F^z in the following way:

$$\Phi(F(\mu)) = \Phi\left(F^Z\left(F^{Z-1}\left(F^{Z-2}\left(...\left(F^2\left(F^1(\mu),\mu\right)\right)..\right)\mu\right)\mu\right)\right)$$
$$:= \left(\Phi \circ F^Z(\mu) \circ F^{Z-1}(\mu) \circ F^{Z-2}(\mu) \circ ... \circ F^2(\mu) \circ F^1\right)(\mu). \tag{24}$$

The sub-functions F^z are defined by the SOR-method that is used to solve the

5 forward model (see Part I). The SOR-method is an iterative approach and the z-th

iteration yields the intermediate result $\Psi_{k,i,j}^z$. The radiance vector Ψ is of length $p =$

$K \times I \times J$ with $k \in [1,K]$, $i \in [1,I]$, and $j \in [1,J]$. The detector readings $P(\mu)$ are the angular-

dependent radiances $\Psi_{k,i,j}^Z$ at the last iteration step Z at detector positions (i,j) on the

boundary. The computational graph of the forward model is depicted in FIG 3A. Starting

10 with the internal properties μ as the input variables, the sub-function F^1 produces the

intermediate result and output variable Ψ^1. The output variables Ψ^z of F^z and the internal

properties μ always serve as input variables for the next sub-function F^{z+1} for all

subsequent steps of the decomposition. The step F^Z calculates the predictions P, which

become the input to the final step of the objective function Φ determination, which is the

15 calculation of the scalar φ.

The sub-functions F^z map the intermediate variables Ψ^{z-1} and constant input

values of μ onto the intermediate result $\Psi^Z = F^Z\left(\Psi^{Z-1}, \mu\right)$ for all iteration steps of the

transport forward model:

$$F^z : \mathfrak{R}^p \times \mathfrak{R}^n \to \mathfrak{R}^p \qquad (25a)$$

$$\begin{pmatrix} \Psi^{z-1} \\ \mu \end{pmatrix} \mapsto \Psi^z, \qquad (25b)$$

For the ordinates with $\xi_k > 0$, $\eta_k > 0$ this mapping is given explicitly by:

5

$$\Psi^z_{k,i,J} = (1-\omega)\Psi^{z-1}_{k,i,J} + \omega \frac{\left\{ S_{k,i,J} + \mu_s \sum_{k=1}^{K} a_k p_{k,k'} \Psi^{z-1}_{k',i,J} + \frac{\xi_k}{\Delta x} \Psi^z_{k,i-1,J} + \frac{\eta_k}{\Delta y} \Psi^z_{k,i,J-1} \right\}}{\left\{ \frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s) \right\}} \qquad (26)$$

For the first iteration step $z=1$, F^z only maps the input variables μ:

$$F^1 : \mathfrak{R}^n \to \mathfrak{R}^p \qquad (27a)$$

10

$$\mu \mapsto \Psi^1 \qquad (27b)$$

and produces:

$$\Psi^1_{k,i,J} = \omega \frac{\left\{ S_{k,i,J} + \frac{\xi_k}{\Delta x} \Psi^1_{k,i-1,J} + \frac{\eta_k}{\Delta y} \Psi^1_{k,i,J-1} \right\}}{\left\{ \frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s) \right\}}. \qquad (28)$$

Equations 26 and 28 are the smallest computational units in the forward code and

play an important role in the adjoint differentiation step, which is discussed below. The

15   gradient $\nabla_\mu \Phi$ of the objective function is obtained by differentiating equation 24 with

respect to the internal properties μ. The total derivative $\nabla_\mu \Phi$ can be obtained by

systematically applying the chain rule of differentiation to equation 22. The total

derivative $\nabla_\mu \Phi$ is a composition of derivatives of all intermediate steps of the forward

21

model due to the explicit dependence of the intermediate results on the internal

properties. Starting from the last step of the forward code, in which the objective function

Φ is calculated given the predicted detector readings $P = \Psi^Z$, the total derivative is given

by the chain rule as:

$$\left(\frac{d\Phi}{d\mu}\right)^T = \left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{d\Psi^Z}{d\mu} + \left(\frac{\partial\Phi}{\partial\mu}\right)^T . \qquad (29)$$

The derivative $\dfrac{d\Psi^Z}{d\mu}$ is obtained by again applying the chain rule of

differentiation:

$$\frac{d\Psi^Z}{d\mu} = \frac{\partial\Psi^Z}{\partial\Psi^{Z-1}}\frac{d\Psi^{Z-1}}{d\mu} + \frac{\partial\Psi^Z}{\partial\mu} . \qquad (30)$$

Equations 29 and 30 yield together:

$$\left(\frac{d\Phi}{d\mu}\right)^T = \left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\Psi^{Z-1}}\frac{d\Psi^{Z-1}}{d\mu} + \left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\mu} + \left(\frac{\partial\Phi}{\partial\mu}\right)^T . \qquad (31)$$

In the next step the total derivative $\dfrac{d\Psi^{Z-1}}{d\mu}$ is replaced again. This process is

repeated for all total derivatives down to the first step $\dfrac{d\Psi^1}{d\mu} = \dfrac{\partial\Psi^1}{\partial\mu}$, to obtain:

$$\left(\frac{d\Phi}{d\mu}\right)^T = \left\{\left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\Psi^{Z-1}}\cdots\frac{\partial\Psi^2}{\partial\Psi^1}\frac{\partial\Psi^1}{\partial\mu}\right\} + \left\{\left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\Psi^{Z-1}}\cdots\frac{\partial\Psi^3}{\partial\Psi^2}\frac{\partial\Psi^2}{\partial\mu}\right\} + \ldots + \left\{\left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\mu}\right\} + \left(\frac{\partial\Phi}{\partial\mu}\right)^T \quad (32)$$

or, using the short hand notation

$$\left\{\left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\Psi^{Z-1}}\cdots\frac{\partial\Psi^{z+1}}{\partial\Psi^z}\right\} = \frac{\partial\left(\Phi\circ\Psi^Z\circ\ldots\circ\Psi^{z+1}\right)^T}{\partial\Psi^z} = \left(\frac{\partial\Phi}{\partial\Psi^z}\right)^T \qquad (33)$$

rewriting equation 33 produces:

597699_1

$$\left(\frac{d\Phi}{d\mu}\right)^T = \left(\frac{\partial\Phi}{\partial\Psi^1}\right)^T \frac{\partial\Psi^1}{\partial\mu} + \left(\frac{\partial\Phi}{\partial\Psi^2}\right)^T \frac{\partial\Psi^2}{\partial\mu} + \dots + \left(\frac{\partial\Phi}{\partial\Psi^z}\right)^T \frac{\partial\Psi^z}{\partial\mu} + \dots + \left(\frac{\partial\Phi}{\partial\Psi^Z}\right)^T \frac{\partial\Psi^Z}{\partial\mu} + \left(\frac{\partial\Phi}{\partial\mu}\right)^T \quad (34)$$

or, $\quad \nabla_\mu \Phi := \left(\frac{d\Phi}{d\mu}\right)^T = \left(\sum_z \left(\frac{\partial\Phi}{\partial\Psi^z}\right)^T \frac{\partial\Psi^z}{\partial\mu}\right) + \left(\frac{\partial\Phi}{\partial\mu}\right)^T, \qquad (35)$

which contains three distinct terms. The last term $\left(\frac{\partial\Phi}{\partial\mu}\right)^T$ is zero, because Φ is not an

explicit function of the internal properties. The middle term $\dfrac{\partial\Psi^z}{\partial\mu}$ can be calculated from

5   equation 26 of the forward model. The derivatives with respect to μ_a and μ_s are:

$$\left(\frac{\partial\Psi^z}{\partial\mu_s}\right)_{k,I,J} = \omega \frac{\sum_{k'=1}^{K} a_k p_{k,k'}\Psi^{z-1}{}_{k',I,J}}{\left\{\frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s)\right\}} - \omega \frac{\left\{S_{k,I,J} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'}\Psi^{z-1}{}_{k',I,J} + \frac{\xi_k}{\Delta x}\Psi^z{}_{k,I-1,J} + \frac{\eta_k}{\Delta y}\Psi^z{}_{k,I,J-1}\right\}}{\left\{\frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s)\right\}^2}$$

$$(36a)$$

$$\left(\frac{\partial\Psi^z}{\partial\mu_a}\right)_{k,I,J} = -\omega \frac{\left\{S_{k,I,J} + \mu_s \sum_{k'=1}^{K} a_{k'} p_{k,k'}\Psi^{z-1}{}_{k',I,J} + \frac{\xi_k}{\Delta x}\Psi^z{}_{k,I-1,J} + \frac{\eta_k}{\Delta y}\Psi^z{}_{k,I,J-1}\right\}}{\left\{\frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s)\right\}^2}. \quad (36b)$$

At this point the adjoint differentiation technique has not yet been used, since the process

10   has not stepped backward through the forward code. This procedure comes into play in

the calculation of the first term $\left(\frac{\partial\Phi}{\partial\Psi^z}\right)^T$ in equation 35. This can be best understood

while looking at the computational graph of the adjoint differentiation technique in FIG

3B. Starting with the last step (output) of the forward code, which is the calculation of

the objective function given the predictions $P = \Psi^Z$, Φ is differentiated with respect to

Ψ^Z and multiply it with $\left(\dfrac{\partial\Phi}{\partial\Phi}\right)^{T}=1$. The result is difference between the measured and

predicted data,

$$\left(\frac{\partial\Phi}{\partial\Psi^{Z}}\right)^{T}=\left(\Psi^{Z}-\mathbf{M}\right)^{T}.\qquad(37)$$

This is the input to our adjoint model, which will eventually provide the gradient

5   $\nabla_{\mu}\Phi$. More specifically, $\left(\dfrac{\partial\Phi}{\partial\Psi^{Z-1}}\right)^{T}$ is calculated by continuing to step backward through

the forward code and is given by:

$$\left(\frac{\partial\Phi}{\partial\Psi^{Z-1}}\right)^{T}=\left(\frac{\partial\Psi^{Z}}{\partial\Psi^{Z-1}}\right)^{T}\left(\frac{\partial\Phi}{\partial\Psi^{Z}}\right)^{T}.\qquad(38)$$

The remaining derivatives $\left(\dfrac{\partial\Phi}{\partial\Psi^{z}}\right)^{T}$ of all intermediate steps in equation 35 are

computed using the previously calculated derivatives $\left(\dfrac{\partial\Phi}{\partial\Psi^{z+1}}\right)^{T}$, such that

$$\begin{aligned}\left(\frac{\partial\Phi}{\partial\Psi^{z}}\right)^{T}&=\frac{\partial\left(\Phi\circ...\circ\Psi^{z+1}\right)\left(\Psi^{z}\right)^{T}}{\partial\Psi^{z}}\\&=\left(\frac{\partial\Psi^{z+1}}{\partial\Psi^{z}}\right)^{T}\cdot\frac{\partial\left(\Phi\circ...\circ\Psi^{z+2}\right)\left(\Psi^{z+1}\right)^{T}}{\partial\Psi^{z+1}}\\&=\left(\frac{\partial\Psi^{z+1}}{\partial\Psi^{z}}\right)^{T}\cdot\left(\frac{\partial\Phi}{\partial\Psi^{z+1}}\right)^{T}\end{aligned}\qquad(39)$$

This step, in which $\left(\dfrac{\partial\Phi}{\partial\Psi^{z}}\right)^{T}$ is calculated from $\left(\dfrac{\partial\Phi}{\partial\Psi^{z+1}}\right)^{T}$, constitutes the adjoint

differentiation step in our code. The matrix $\left(\dfrac{\partial\Psi^{z+1}}{\partial\Psi^{z}}\right)^{T}$ is calculated by analytically

24

597699_1

differentiating the $(z+1)$-th SOR-iteration step, given in equation 26, with respect to Ψ.

We get, for example in the case of ordinates with $\xi_k > 0, \eta_k > 0$:

$$\frac{\partial \Psi^{z+1}_{k,I,J}}{\partial \Psi^z_{k,i',J'}} = (1-\omega)\delta_{k,I,J} + \omega \frac{\left\{ \mu_s a_k p_{k,k} \delta_{I,J} + \frac{\xi_k}{\Delta x} \delta_{k,I-1,J} + \frac{\eta_k}{\Delta y} \delta_{k,I,J-1} \right\}}{\left\{ \frac{\xi_k}{\Delta x} + \frac{\eta_k}{\Delta y} + (\mu_a + \mu_s) \right\}} \qquad (40)$$

with $\delta_{a,b,c} = \delta_a \delta_b \delta_c$ and $\delta_x = \begin{cases} 1 & if \quad x = x' \\ 0 & if \quad x \neq x' \end{cases}$.

5　　　Here, the approximations $\frac{\xi_k}{\Delta x} \frac{\partial \Psi^{z+1}_{k,I-1,J}}{\partial \Psi^z_{k,I,J}} \cong \frac{\xi_k}{\Delta x} \delta_{k,I-1,J}$ and $\frac{\eta_k}{\Delta y} \frac{\partial \Psi^{z+1}_{k,I,J-1}}{\partial \Psi^z_{k,I,J}} \cong \frac{\eta_k}{\Delta y} \delta_{k,I,J-1}$

can be made for the relevant terms on the right hand side of equation 40.

As can be seen, the gradient of the objective function is calculated by stepping

backward through all previously calculated iteration steps of the forward model without

solving an entirely new numerical adjoint equation of radiative transfer. Furthermore, the

10　　particular underlying physical system does not have to be known, because the derivative

is computed directly from the elementary steps of the forward model code (equations 26

and 28).

3.　　Gradient-Based Optimization

Once the gradient $\nabla_\mu \Phi(\mu_0)$ for a point μ_0 is obtained, a one-dimensional line

15　　minimization along the given gradient direction is performed until a minimum, $\Phi(\mu_{min})$ is

found, FIG. 1, step 135 (inner iteration). Various techniques can be applied to perform

such one-dimensional minimizations. Exemplary techniques are disclosed in Nash S G,

Sofer A. Linear and nonlinear programming, McGraw-Hill, New York, 1996, and Press

W H, Teukolsky S A, Vetterling W T, and Flannery B P. Numerical Recipes in C.

Cambridge University Press, New York, (1992). One approach is to start with three points, the initial guess μ_0, $\mu_1 = \mu_0 + \Delta\mu$, and $\mu_2 = \mu_0 + 2\Delta\mu$ chosen along the direction of the gradient. After calculating the objective function at these three points, the largest calculated value is neglected, and a new guess μ_3 is determined using the golden section

5   rule until a minimum is bracketed. At that time a parabola is fit through these three points. The smallest value of the objective function on this parabola is determined to be the minimum. Once the point (guess) $\mu_{min,1}$ is found for which $\Phi(\mu_{min,1})$ is minimal along the one-dimensional subspace, a new gradient calculation $\nabla_\mu\Phi(\mu_{min,1})$ is performed at the minimum $\mu_{min,1}$ and a new direction is chosen in a conjugate-gradient framework. For

10   this a Polak-Riberie conjugate-gradient scheme can be employed (outer iterations.)

After completing several inner and outer iterations a final point $\mu_{min,final}$ is found for which $\Phi(\mu_{min,final})$ is smaller than for all other points μ_{min} before. The final reconstruction image is than given by $\mu_{min,final}$, wherein $\mu_{min,final}$ represents the spatial distribution of the optical properties in the target medium. The image can be represented

15   by any known means, such as on a computer monitor or printed as a hard copy on paper, wherein the property values of the medium may be represented as shades of gray or colors.

Although illustrative embodiments have been described herein in detail, those skilled in the art will appreciate that variations may be made without departing from the

20   spirit and scope of this invention. Moreover, unless otherwise specifically stated, the terms and expressions used herein are terms of description and not terms of limitation, and are not intended to exclude any equivalents of the system and methods set forth in the following claims.

26